On the Extension of Linear Contractions: Some Results of the Hahn-Banach Type

W. Hässelbarth
Bundesanstalt für Materialforschung und -prüfung (BAM)
Berlin, Germany

and

E. Ruch
Max-Planck-Institut für Biophysikalische Chemie
Göttingen, Germany

Submitted by T. Ando

ABSTRACT

Let $\|\cdot\|$ be a norm on R^n, and $\mathscr{S}$ a compact convex semigroup of linear $\|\cdot\|$-contractions. Given two k-tuples of n-vectors, $(x^{(1)},\dots,x^{(k)})$ and $(y^{(1)},\dots,y^{(k)})$, we seek conditions for the existence of a contraction $S\in\mathscr{S}$ that simultaneously takes $x^{(i)}$ to $y^{(i)}$, that is, $Sx^{(i)}=y^{(i)}$, for all $i=1,\dots,k$!. Straightforward application of the separation theorem for convex sets provides a general but abstract result, in the form of a system of inequalities. Specializing $\|\cdot\|$ to the 1-norm and the ∞-norm, respectively, and $\mathscr{S}$ to comprise either all contractions or those contractions that preserve a particular linear form, it is possible to evaluate the characteristic functionals arising from the separation theorem. Thereby the abstract result can be reduced to a tractable form, which turns out to be of the Hahn-Banach type.

INTRODUCTION

Recently Schranner, Seligman, and one of the authors [8] proved

THEOREM A. *The following assertions about pairs (x, y) and (x', y') of n-dimensional probability vectors are equivalent*:

LINEAR ALGEBRA AND ITS APPLICATIONS 184: 145–163 (1993)

655 Avenue of the Americas, New York, NY 10010 0024-3795/93/$6.00

(I) *There is a stochastic matrix S which simultaneously takes x to x' and y to y', that is,* $Sx = x'$ *and* $Sy = y'$.

(II) *The mixing distance of the pair* (x, y) *is greater than or equal to that of the pair* (x', y'), *that is,*

$$\|x' - \tau y'\|_1 \leqslant \|x - \tau y\|_1 \qquad \textit{for all} \quad 0 \leqslant \tau \in R.$$

Here, as usual, the 1-*norm* of a vector $x = [x_i] \in R^n$ is the sum of the absolute values of its components, $\|x\|_1 = \Sigma|x_i|$. A matrix is *stochastic* if all of its entries are nonnegative and all column sums are equal to one, and a vector is called a *probability vector* if all of its components are nonnegative and add up to one.

With the particular choice $y' = y = e_0$, the constant vector with components whose sum is equal to one, this reduces to a well-known theorem due to Hardy, Littlewood, and Pólya (see [4]);

THEOREM B. *The following statements about two probability vectors x and x' are equivalent*:

(I) *There is a doubly stochastic matrix D such that* $Dx = x'$.

(II) *The mixing character of* x' *is greater than or equal to that of* x, *that is,*

$$\|x' - \tau e_0\|_1 \leqslant \|x - \tau e_0\|_1 \qquad \textit{for all} \quad 0 \leqslant \tau \in R.$$

Here a matrix is *doubly stochastic* if it is stochastic and moreover all row sums are equal to one.

Remark that the implication (I) $\Rightarrow$ (II) is trivial in each of Theorems A and B, and that the inequality in (II) is true for all $\tau \in R$.

Observing that a stochastic matrix is characterized as a (linear) trace-preserving $\|\cdot\|_1$-contraction of R^n and that stochastic matrices form a compact convex (multiplicative) semigroup with identity, in this paper we treat the following problem: Let $\|\cdot\|$ be a norm on R^n, and $\mathscr{S}$ be a compact convex semigroup with identity of contractions of $(R^n, \|\cdot\|)$. Given two k-tuples of vectors $[x^{(j)}]$, $[y^{(j)}]$ of R^n, find conditions for the existence of a contraction $S \in \mathscr{S}$ that simultaneously takes $x^{(j)}$ to $y^{(j)}$, that is, $Sx^{(j)} = y^{(j)}$ for all $j = 1, \ldots, k$.

On the basis of the familiar separation theorem for compact convex subsets, it is not difficult to give an answer to this problem in an abstract form (Theorem 1). We are especially interested in the case when $\mathscr{S}$ consists of *all* (linear) $\|\cdot\|$-contractions, and in particular, when $\|\cdot\|$ is specialized to the 1-norm and the ∞-norm, respectively.

When specialized to the ∞-norm, $\|x\|_\infty = \max_i |x_i|$ for $x = [x_i]$, the theorem says that there exists a $\|\cdot\|_\infty$-contraction of R^n that simultaneously takes $x^{(j)}$ into $y^{(j)}$ for all $j = 1, \dots, k$ if and only if

$$\left\| \sum_{j=1}^{k} \alpha_j y^{(j)} \right\|_\infty \leqslant \left\| \sum_{j=1}^{k} \alpha_j x^{(j)} \right\|_\infty \qquad \text{for all} \quad \alpha_1, \dots, \alpha_k \in R$$

(Theorem 2). Since the above inequality implies that the correspondence $\sum \alpha_j x^{(j)} \to \sum \alpha_j y^{(j)}$ defines a $\|\cdot\|_\infty$-contractive linear map from the subspace of R^n spanned by the $x^{(j)}$ ($j = 1, \dots, k$) into R^n, Theorem 2 is equivalent to stating that any $\|\cdot\|_\infty$-contractive linear map from a subspace of R^n into R^n can be extended to a $\|\cdot\|_\infty$-contraction of R^n (Corollary 3). In this form, however, the assertion is a corollary of the classical Kantorovich theorem, a generalization of the Hahn-Banach theorem.

In the other special case considered here, that of the 1-norm, the system of inequalities according to Theorem 1 cannot be reduced to a convenient and simple form like that in Theorem 2, valid for all $k < n$. For $k = 2$, however, owing to a special property of polyhedral norms on R^2 (Lemma 4), as the key result of this paper, an analogue of Theorem 2 can be proved: There exists a $\|\cdot\|_1$-contraction of R^n that simultaneously takes $x^{(j)}$ to $y^{(j)}$ for $j = 1, 2$ if and only if

$$\|\alpha_1 y^{(1)} + \alpha_2 y^{(2)}\|_1 \leqslant \|\alpha_1 x^{(1)} + \alpha_2 x^{(2)}\|_1 \qquad \text{for all} \quad \alpha_1 \in R, \quad \alpha_2 \in R$$

(Theorem 5). This result generalizes Theorem A from trace-preserving $\|\cdot\|_1$-contractions to the case of general $\|\cdot\|_1$-contractions.

Theorem 5 can also be reformulated as a Hahn-Banach type result as follows (Corollary 6): Any $\|\cdot\|_1$-contractive linear map into R^n, defined on a subspace of R^n of dimension $\leqslant 2$, can be extended to a $\|\cdot\|_1$-contraction of R^n. In addition, it can be shown that this is not true, in general, for $\|\cdot\|_1$-contractions defined on subspaces of dimension $\geqslant 3$.

Related problems have been treated by Alberti and Uhlmann [1], using techniques based on separation theorems for convex sets. In this paper we use the same approach as the principal theoretical tool, but then put the main effort into reduction of the abstract results to a tractable form. The interpretation of our results as generalizations of the Hahn-Banach theorem has been motivated by the presentation of results from [1] and [8] in a recent paper of Ando [2].

1. ABSTRACT FORM

Let $\mathcal{S}$ be a compact convex (multiplicative) semigroup with identity of linear maps of R^n. For a vector $x \in R^n$, denote by $\mathcal{S}(x)$ the *$\mathcal{S}$-orbit* of x, $\mathcal{S}(x) = \{x' = Sx;\ S \in \mathcal{S}\}$. Then for a vector $y \in R^n$, there exists an $S \in \mathcal{S}$ that takes x to y if and only if y belongs to $\mathcal{S}(x)$. Since by assumption $\mathcal{S}$ is a compact convex set of operators, $\mathcal{S}(x)$ is a compact convex subset of R^n. Since every linear form on R^n is given by the scalar product $\langle z, \cdot \rangle$ with a vector $z \in R^n$, the separation theorem for convex sets (see [5]) implies that $y \in \mathcal{S}(x)$ if and only if

$$\langle z, y \rangle \leqslant \sup_{S \in \mathcal{S}} \langle z, Sx \rangle \qquad \text{for all} \quad z \in R^n. \tag{1.1}$$

For each $z \in R^n$, define a functional $s_z(\cdot)$ of R^n by

$$s_z(w) = \sup_{S \in \mathcal{S}} \langle z, Sw \rangle \qquad \text{for all} \quad w \in R^n. \tag{1.2}$$

Since by assumption $\mathcal{S}$ is a multiplicative semigroup with identity, $y \in \mathcal{S}(x)$ is equivalent to $\mathcal{S}(y) \subset \mathcal{S}(x)$. Therefore we can conclude that there is an $S \in \mathcal{S}$ that takes x to y if and only if

$$s_z(y) \leqslant s_z(x) \qquad \text{for all} \quad z \in R^n. \tag{1.3}$$

Now consider the direct sum $R^{kn} = R^n \oplus R^n \oplus \cdots \oplus R^n$ of k copies of R^n, consisting of k-tuples $w = (w^{(1)}, \ldots, w^{(k)})$ of vectors $w^{(j)} \in R^n$. Remark that every linear form $f(\cdot)$ on R^{kn} is given by the scalar product (inherited from that of R^n) with a vector $z = (z^{(1)}, \ldots, z^{(k)})$, that is,

$$f(w) = \sum_{j=1}^{k} \langle z^{(j)}, w^{(j)} \rangle \qquad \text{for all} \quad w \in R^{kn}. \tag{1.4}$$

Consider further the subset $\operatorname{diag} \mathcal{S}^k = \{(S, \ldots, S);\ S \in \mathcal{S}\}$ of linear maps of R^{kn}, defined by the diagonal action of $\mathcal{S}$ on R^{kn}. It is immediate to see that $\operatorname{diag} \mathcal{S}^k$ is a compact convex (multiplicative) semigroup with identity of linear maps of R^{kn}.

Putting together this observation and the general existence criterion stated before leads to the following assertion.

THEOREM 1. *Let $\mathscr{S}$ be a compact convex (multiplicative) semigroup with identity of linear maps of R^n. Given two k-tuples $[x^{(j)}], [y^{(j)}]$ of vectors of R^n, there exists an $S \in \mathscr{S}$ that simultaneously takes $x^{(j)}$ to $y^{(j)}$ for all $j = 1, \ldots, k$ if and only if*

$$s_z(y) \leqslant s_z(x) \qquad \textit{for all} \quad z \in R^{kn}.$$

Here $x = (x^{(1)}, \ldots, x^{(k)})$, $y = (y^{(1)}, \ldots, y^{(k)})$, $z = (z^{(1)}, \ldots, z^{(k)})$, and

$$s_z(x) = \sup_{S \in \mathscr{S}} \sum_{j=1}^{k} \langle z^{(j)}, Sx^{(j)} \rangle. \tag{1.5}$$

2. RESULTS FOR THE ∞-NORM

In this section $\mathscr{S}$ is chosen as a compact convex (multiplicative) semigroup with identity of linear $\|\cdot\|_\infty$-contractions of R^n. Recall that a matrix $S = [S_{pq}]$ gives rise to a $\|\cdot\|_\infty$-contraction if and only if

$$\sum_{q=1}^{n} |S_{pq}| \leqslant 1 \qquad \text{for all} \quad p = 1, \ldots, n. \tag{2.1}$$

A matrix S with this property, in addition, preserves the first component, that is, $(Sw)_1 = w_1$ for $w \in R^n$, if and only if $S_{11} = 1$.

If $\mathscr{S}$ is chosen as the class of all $\|\cdot\|_\infty$-contractions, the functionals $s_z(\cdot)$ occurring in Theorem 1 are readily evaluated as follows: For $x = (x^{(1)}, \ldots, x^{(k)})$ and $z = (z^{(1)}, \ldots, z^{(k)}) \in R^{kn}$, and with the notation $x_i^{(j)}$ for the ith component of the jth vector $x^{(j)} \in R^n$, we have

$$s_z(x) = \sup_{S \in \mathscr{S}} \sum_{j=1}^{k} \langle z^{(j)}, Sx^{(j)} \rangle = \sum_{p=1}^{n} \max_{1 \leqslant q \leqslant n} \left| \sum_{j=1}^{k} z_p^{(j)} x_q^{(j)} \right|$$

$$= \sum_{p=1}^{n} \left\| \sum_{j=1}^{k} z_p^{(j)} x^{(j)} \right\|_\infty . \tag{2.2}$$

If $\mathscr{S}$ is chosen as the class of $\|\cdot\|_\infty$-contractions which preserve the first component, we similarly arrive at

$$s_z(x) = \sum_{p=2}^{n} \left\| \sum_{j=1}^{k} z_p^{(j)} x^{(j)} \right\|_\infty + \sum_{j=1}^{k} z_1^{(j)} x_1^{(j)}. \tag{2.3}$$

Since the $[z_p^{(j)}]$ can be arbitrarily chosen, we see that when $\mathscr{S}$ is the class of all $\|\cdot\|_\infty$-contractions,

$$s_z(y) \leqslant s_z(x) \qquad \text{for all} \quad z \in R^{kn} \tag{2.5}$$

holds if and only if

$$\left\| \sum_{j=1}^{k} \alpha_j y^{(j)} \right\|_\infty \leqslant \left\| \sum_{j=1}^{k} \alpha_j x^{(j)} \right\|_\infty \qquad \text{for all} \quad \alpha_j \in R, \quad j = 1, \dots, k. \tag{2.6}$$

Similarly, when $\mathscr{S}$ is the class of $\|\cdot\|_\infty$-contractions preserving the first component, (2.5) is valid if and only if, in addition to (2.6), the following relations are satisfied:

$$y_1^{(j)} = x_1^{(j)} \qquad \text{for all} \quad j = 1, \dots, k. \tag{2.7}$$

Summing up, we can see that Theorem 1 takes the following forms for those classes of $\|\cdot\|_\infty$-contractions considered above.

THEOREM 2. *Let $[x^{(j)}]$ and $[y^{(j)}]$ be two k-tuples of vectors of R^n.*

(I) *There is a $\|\cdot\|_\infty$-contraction that simultaneously takes $x^{(j)}$ to $y^{(j)}$ for $j = 1, \dots, k$ if and only if* (2.6) *is satisfied, that is,*

$$\left\| \sum_{j=1}^{k} \alpha_j y^{(j)} \right\|_\infty \leqslant \left\| \sum_{j=1}^{k} \alpha_j x^{(j)} \right\|_\infty \qquad \textit{for all} \quad \alpha_j \in R, \quad j = 1, \dots, k.$$

(II) *There is a* $\|\cdot\|_\infty$*-contraction preserving the first component that simultaneously takes* $x^{(j)}$ *to* $y^{(j)}$ *for* $j = 1, \dots, k$ *if and only if both* (2.6) *and* (2.7) *are satisfied, that is,*

$$\left\| \sum_{j=1}^{k} \alpha_j y^{(j)} \right\|_\infty \leqslant \left\| \sum_{j=1}^{k} \alpha_j x^{(j)} \right\|_\infty \qquad \textit{for all} \quad \alpha_j \in R, \quad j = 1, \dots, k,$$

and

$$y_1^{(j)} = x_1^{(j)} \qquad \textit{for all} \quad j = 1, \dots, k.$$

Now suppose that there is given a $\|\cdot\|_\infty$-contractive linear map T from the subspace of R^n spanned by $x^{(1)}, \dots, x^{(k)}$ into R^n. Let $y^{(j)} = Tx^{(j)}$ for $j = 1, \dots, k$. Then the $\|\cdot\|_\infty$-contractivity of T implies that

$$\left\| \sum_{j=1}^{k} \alpha_j y^{(j)} \right\|_\infty = \left\| T\left(\sum_{j=1}^{k} \alpha_j x^{(j)} \right) \right\|_\infty \leqslant \left\| \sum_{j=1}^{k} \alpha_j x^{(j)} \right\|_\infty \qquad \text{for all} \quad \alpha_j \in R,$$

$$j = 1, \dots, k. \tag{2.8}$$

Therefore (2.6) is satisfied for the k-tuples $[x^{(j)}]$ and $[y^{(j)}]$. Then by Theorem 2 there is a $\|\cdot\|_\infty$-contraction S of R^n such that

$$Tx^{(j)} = Sx^{(j)} \qquad \text{for all} \quad j = 1, \dots, k, \tag{2.9}$$

that is, S is a contractive extension of T to the whole space R^n. If in addition T preserves the first component, that is,

$$(Tx^{(j)})_1 = x_1^{(j)} \qquad \text{for all} \quad j = 1, \dots, k, \tag{2.10}$$

then T can be extended to a $\|\cdot\|_\infty$-contraction which also preserves the first component.

Therefore Theorem 2 can be equivalently expressed in the following form.

COROLLARY 3. *Any* $\|\cdot\|_\infty$*-contractive linear map* T *from a subspace of* R^n *into* R^n *can be extended to a linear* $\|\cdot\|_\infty$*-contraction* S *of* R^n. *If, in addition,* T *preserves the first component, so does* S.

The first part of Corollary 3, in fact, is a special case of a classical theorem, which is usually attributed to Kantorovich (see e.g. [6]) but appears to go back to Nachbin [7]. This theorem states that any contractive linear map from a subspace of a Banach space X into a Banach space Y of L^∞ type can be extended to a linear contraction from X to Y. The second part of Corollary 3 can also be derived from this theorem.

We should like to emphasize that Corollary 3 does not hold trivially, as would be the case if any subspace X of R^n had a contractive projector P onto it. Then, if T is a contractive linear map from X to R^n, trivially, $S = TP$ would be a contractive linear map on R^n, extending T. This is e.g. true in the case of the euclidean metric, that is, for the 2-norm of R^n, but it is not so for the ∞-norm.

3. RESULTS FOR THE 1-NORM

In this section $\mathcal{S}$ is chosen as a compact convex (multiplicative) semigroup with identity of linear $\|\cdot\|_1$-contractions of R^n. Recall that a matrix $S = [S_{pq}]$ gives rise to a $\|\cdot\|_1$-contraction if and only if

$$\sum_{p=1}^{n} |S_{pq}| \leqslant 1 \qquad \text{for all} \quad q = 1, \ldots, n. \tag{3.1}$$

A matrix with this property in addition preserves the trace—that is, $\Sigma(Sw)_i = \operatorname{tr}(Sw) = \operatorname{tr}(w) = \Sigma w_i$ for all $w \in R^n$—if and only if

$$\sum_{p=1}^{n} S_{pq} = 1 \qquad \text{for all} \quad q = 1, \ldots, n. \tag{3.2}$$

Note that a trace-preserving linear $\|\cdot\|_1$-contraction is nothing but a stochastic map.

If $\mathcal{S}$ is chosen as the class of all $\|\cdot\|_1$-contractions, the functionals $s_z(\cdot)$ defined in (1.5) are evaluated as follows. For $x = (x^{(1)}, \ldots, x^{(k)})$ and $z = (z^{(1)}, \ldots, z^{(k)}) \in R^{kn}$, and with $x_i^{(j)}$ denoting the ith entry in the jth vector $x^{(j)} \in R^n$, as in the preceding section, we have

$$s_z(x) = \sup_{S \in \mathcal{S}} \sum_{j=1}^{k} \langle z^{(j)}, Sx^{(j)} \rangle = \sum_{q=1}^{n} \max_{1 \leqslant p \leqslant n} \left| \sum_{j=1}^{k} z_p^{(j)} x_q^{(j)} \right|. \tag{3.3}$$

Any k-tuple $u = (u^{(1)}, \ldots, u^{(k)})$ of n-vectors $u^{(j)} \in R^n$ is associated with an n-tuple $u^T = (u_{(1)}, \ldots, u_{(n)})$ of k-vectors $u_{(p)} \in R^k$, whose jth entry is given by $u_p^{(j)}$, for $j = 1, \ldots, k$ and $p = 1, \ldots, n$. Making use of this, the expression for $s_{\tilde{z}}(x)$ can be rewritten as follows:

$$s_{\tilde{z}}(x) = \sum_{q=1}^{n} \max_{1 \leqslant p \leqslant n} |\langle z_{(p)}, x_{(q)} \rangle|. \tag{3.4}$$

Now by Theorem 1, given two k-tuples $[x^{(j)}]$ and $[y^{(j)}]$ of vectors of R^n, there is a $\|\cdot\|_1$-contraction that simultaneously takes $x^{(j)}$ to $y^{(j)}$ for all $j = 1, \ldots, k$ if and only if

$$\sum_{q=1}^{n} \max_{1 \leqslant p \leqslant n} |\langle z_{(p)}, y_{(q)} \rangle| \leqslant \sum_{q=1}^{n} \max_{1 \leqslant p \leqslant n} |\langle z_{(p)}, x_{(q)} \rangle| \tag{3.5}$$

holds for all $(z_{(1)}, \ldots, z_{(n)}) \in R^{nk} = R^k \oplus \cdots \oplus R^k$.

Before trying to simplify these results, let us first turn to the case where $\mathscr{S}$ is the class of linear $\|\cdot\|_1$-contractions preserving the trace, that is, the class of stochastic linear maps of R^n. It is easy to see that in this case

$$s_{\tilde{z}}(x) = \sum_{q=1}^{n} \max_{1 \leqslant p \leqslant n} \langle z_{(p)}, x_{(q)} \rangle. \tag{3.6}$$

Therefore, by Theorem 1, given two k-tuples of vectors $[x^{(j)}]$ and $[y^{(j)}] \in R^n$, there is a stochastic map that simultaneously takes $x^{(j)}$ to $y^{(j)}$ for all $j = 1, \ldots, n$ if and only if

$$\sum_{q=1}^{n} \max_{1 \leqslant p \leqslant n} \langle z_{(p)}, y_{(q)} \rangle \leqslant \sum_{q=1}^{n} \max_{1 \leqslant p \leqslant n} \langle z_{(p)}, x_{(q)} \rangle \tag{3.7}$$

holds for all $z_{(1)}, \ldots, z_{(n)} \in R^k$.

It is obvious that if there is a stochastic map that simultaneously takes $x^{(j)}$ to $y^{(j)}$ for $j = 1, \ldots, k$, then (3.5) is satisfied and, in addition,

$$\operatorname{tr}(y^{(j)}) = \operatorname{tr}(x^{(j)}) \qquad \text{for all} \quad j = 1, \ldots, k. \tag{3.8}$$

Hence (3.7) implies (3.5) and (3.8). The converse implication is not valid in general. It is, however, true if the subspace of R^n spanned by the $x^{(j)}$, $j = 1, \ldots, k$, has a nontrivial intersection with the positive cone of R^n.

Suppose that (3.5) and (3.8) are satisfied, and that some linear combination of the $x^{(j)}$, $j = 1, \dots, k$, has positive entries only. Then, given a k-tuple $[z^{(j)}]$ of vectors of R^n, by assumption there are coefficients $\beta_j \in R$ ($j = 1, \dots, k$) such that

$$\sum_{j=1}^{k} z_p^{(j)} x_q^{(j)} + \sum_{j=1}^{k} \beta_j x_q^{(j)} > 0 \qquad \text{for all} \quad p, q = 1, \dots, n. \tag{3.9}$$

For $j = 1, \dots, n$, let $w^{(j)} = z^{(j)} + \beta_j e$, where $e \in R^n$ is the constant vector with entries 1. Then

$$\sum_{q=1}^{n} \max_{1 \leqslant p \leqslant n} |\langle w_{(p)}, x_{(q)} \rangle| = \sum_{q=1}^{n} \max_{1 \leqslant p \leqslant n} \langle z_{(p)}, x_{(q)} \rangle + \sum_{j=1}^{k} \beta_j \operatorname{tr}(x^{(j)}).$$

On the other hand, by (3.5)

$$\begin{aligned} \sum_{q=1}^{n} \max_{1 \leqslant p \leqslant n} |\langle w_{(p)}, x_{(q)} \rangle| &\geqslant \sum_{q=1}^{n} \max_{1 \leqslant p \leqslant n} |\langle w_{(p)}, y_{(q)} \rangle| \\ &\geqslant \sum_{q=1}^{n} \max_{1 \leqslant p \leqslant n} \langle w_{(p)}, y_{(q)} \rangle \\ &= \sum_{q=1}^{n} \max_{1 \leqslant p \leqslant n} \langle z_{(p)}, y_{(q)} \rangle + \sum_{j=1}^{k} \beta_j \operatorname{tr}(y^{(j)}). \end{aligned}$$

Hence by the trace preservation condition (3.8) we conclude that (3.7) is satisfied.

Summarizing these results, if some linear combination of the $x^{(j)}$, $j = 1, \dots, k$, has positive entries only, there is a stochastic linear map that takes $x^{(j)}$ to $y^{(j)}$ for $j = 1, \dots, k$ if and only if (3.5) and the trace preservation condition (3.8) are satisfied.

The expressions (3.4) and (3.6) for the functionals $s_{\underset{\sim}{z}}(\cdot)$ and the corresponding systems of inequalities (3.5) and (3.7) cannot be reduced to an essentially simpler form, expressed in terms of 1-norms, in general. However, for $k = 2$, that is, if $z_{(p)}$ and $x_{(p)}$ are 2-vectors, such reduction is possible on the basis of the following lemma.

LEMMA 4.

(1) *For any finite set* $\{z_{(1)}, \ldots, z_{(n)}\}$ *of vectors* $z_{(p)} \in R^2$ *there is a finite set* $\{w_{(1)}, \ldots, w_{(n)}\}$ *of vectors* $w_{(p)} \in R^2$ *such that*

$$\max_{1 \leqslant p \leqslant n} |\langle z_{(p)}, x \rangle| = \sum_{p=1}^{n} |\langle w_{(p)}, x \rangle| \qquad \textit{for all} \quad x \in R^2.$$

(2) *Conversely, for any finite set* $\{w_{(1)}, \ldots, w_{(n)}\} \subset R^2$ *there is a finite set* $\{z_{(1)}, \ldots, z_{(n)}\} \subset R^2$ *such that*

$$\sum_{p=1}^{n} |\langle w_{(p)}, x \rangle| = \max_{1 \leqslant p \leqslant n} |\langle z_{(p)}, x \rangle| \qquad \textit{for all} \quad x \in R^2.$$

This lemma is based on a special property of polyhedral norms in R^2 and cannot be extended to higher dimensions. A proof will be given in the next section.

Returning to the system of inequalities (3.5), for $k = 2$, we may now by Lemma 4 replace the terms $\max|\langle z_{(p)}, x_{(q)} \rangle|$ with $\Sigma|\langle w_{(p)}, x_{(q)} \rangle|$. Letting $w^{(1)}$ and $w^{(2)}$ be the n-vectors whose entries are the first components of $w_{(1)}, \ldots, w_{(n)}$ and the second components of $w_{(1)}, \ldots, w_{(n)}$, respectively, we have

$$s_z(x) = \sum_{q=1}^{n} \sum_{p=1}^{n} |\langle w_{(p)}, x_{(q)} \rangle| = \sum_{p=1}^{n} \left\| \sum_{j=1}^{2} w_p^{(j)} x^{(j)} \right\|_1. \tag{3.10}$$

Noting that by Lemma 4 $(w^{(1)}, w^{(2)})$ runs over all 2-tuples of n-vectors if $(z^{(1)}, z^{(2)})$ does, we can conclude that, for $k = 2$, the condition (3.5) is valid if and only if

$$\left\| \sum_{j=1}^{2} \alpha_j y^{(j)} \right\|_1 \leqslant \left\| \sum_{j=1}^{2} \alpha_j x^{(j)} \right\|_1 \qquad \text{for all} \quad \alpha_j \in R, \quad j = 1, 2. \tag{3.11}$$

Summing up, we have the following theorem.

THEOREM 5.

(I) *Given four vectors* $x^{(1)}, x^{(2)}, y^{(1)}, y^{(2)} \in R^n$, *there is a linear* $\|\cdot\|_1$-*contraction of* R^n *that simultaneously takes* $x^{(1)}$ *to* $y^{(1)}$ *and* $x^{(2)}$ *to* $y^{(2)}$ *if and only if* (3.11) *is satisfied, that is,*

$$\|\alpha_1 y^{(1)} + \alpha_2 y^{(2)}\|_1 \leqslant \|\alpha_1 x^{(1)} + \alpha_2 x^{(2)}\|_1 \qquad \textit{for all} \quad \alpha_1, \alpha_2 \in R.$$

(II) *Given four vectors* $x^{(1)}$, $x^{(2)}$, $y^{(1)}$, $y^{(2)} \in R^n$ *such that some linear combination of* $x^{(1)}$ *and* $x^{(2)}$ *has positive entries only, there is a stochastic linear map of* R^n *that simultaneously takes* $x^{(1)}$ *to* $y^{(1)}$ *and* $x^{(2)}$ *to* $y^{(2)}$ *if and only if the condition above is satisfied and, in addition,*

$$\operatorname{tr}(y^{(j)}) = \operatorname{tr}(x^{(j)}) \qquad \textit{for} \quad j = 1, 2.$$

As in the preceding section, Theorem 5 admits the following equivalent expression.

COROLLARY 6.

(I) *Any* $\|\cdot\|_1$*-contractive linear map from a subspace of* R^n *of dimension* $\leqslant 2$ *into* R^n *can be extended to a linear* $\|\cdot\|_1$*-contraction of* R^n.

(II) *Any trace-preserving linear* $\|\cdot\|_1$*-contraction from a subspace of* R^n *of dimension* $\leqslant 2$, *containing a vector with positive entries only, into* R^n *can be extended to a stochastic linear map of* R^n.

That Corollary 6 can be no longer true for the case of a subspace of dimension 3 can be concluded from an example due to Sherman [9].

4. PROOF OF LEMMA 4

Recall that a norm of R^2 is called *polyhedral* if its unit ball is a polyhedron, that is, the convex hull of a finite number of points. Since the unit ball of a norm is symmetric with respect to the inversion at the origin, it has an even number of extreme points, $x^{(1)}, x^{(2)}, \dots, x^{(2n)}$, which can be arranged in such a way (by anticlockwise ordering) that

$$\begin{aligned} x^{(n+i)} &= -x^{(i)} && \text{for} \quad i = 1, \dots, n, \\ \langle x^{(j)}, Qx^{(k)} \rangle &\geqslant 0 && \text{for} \quad 0 \leqslant k - j \leqslant n. \end{aligned} \tag{4.1}$$

Here Q is the linear (orthogonal) map of R^2 that gives rise to the clockwise rotation by $\pi/2$. Figure 1 shows an example.

Note that for any inversion-symmetric finite convex subset of R^2 the extreme points can be arranged so that (4.1) is valid.

Now given a finite (spanning) subset $A = \{z^{(1)}, \dots, z^{(n)}\}$ of R^2, consider the *max norm* afforded by A,

$$\mu_A(x) = \max_{1 \leqslant i \leqslant n} |\langle z^{(i)}, x \rangle| \qquad \text{for} \quad x \in R^2. \tag{4.2}$$

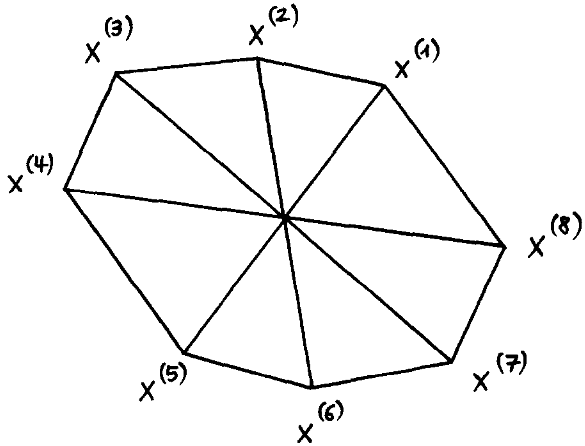

FIG. 1.

Letting A_0 be the set of extreme points of the convex polyhedron spanned by the $2n$ vectors $z^{(i)}$ and $-z^{(i)}$, $i = 1, \ldots, n$, it is obvious that A_0 affords the same max norm as A does.

Letting $z^{(n+i)} = -z^{(i)}$ for $i = 1, \ldots, n$, we may therefore assume that each $z^{(j)}$ ($j = 1, \ldots, 2n$) is an extreme point of the convex hull $\operatorname{conv}\{z^{(j)}; j = 1, \ldots, 2n\}$, and that (4.1) is satisfied with z's in place of x's. For convenience let $z^{(2n+j)} = z^{(j)}$.

The unit ball of the max norm $\mu_A(\cdot)$ is determined by the set of inequalities

$$\langle z^{(j)}, y \rangle \leqslant 1 \qquad \text{for} \quad j = 1, \ldots, 2n, \tag{4.3}$$

and hence is a polyhedron with $2n$ corners, say $x^{(k)}$ ($k = 1, \ldots, 2n$). We may assume that (4.1) is satisfied for the x's.

Letting again $x^{(2n+k)} = x^{(k)}$ for $k = 1, \ldots, 2n$, consider the half space H_k that is bounded by the straight line passing through $x^{(k)}$ and $x^{(k+1)}$, and contains the origin. Letting

$$H_k = \{y \in R^2; \langle w^{(k)}, y \rangle \leqslant 1\}, \tag{4.4}$$

the vector $w^{(k)}$ is uniquely determined by the fact that $x^{(k)}$ and $x^{(k+1)}$ are on the boundary of H^k, that is,

$$\langle w^{(k)}, x^{(k)} \rangle = 1 = \langle w^{(k)}, x^{(k+1)} \rangle, \tag{4.5}$$

or, equivalently,

$$w^{(k)} = \frac{Q(x^{(k+1)} - x^{(k)})}{\langle x^{(k)}, Qx^{(k+1)} \rangle}. \tag{4.6}$$

Here Q again is the orthogonal map of R^2 effecting the clockwise rotation by 90 degrees, which entails the relations

$$\langle x, Qx \rangle = 0 \text{ and } \langle x, Qy \rangle = -\langle y, Qx \rangle \qquad \text{for all} \quad x, y \in R^2. \tag{4.7}$$

Summarizing, we have two representations of the unit ball of μ_A as an intersection of $2n$ half spaces: the half spaces H_k ($k = 1, \ldots, 2n$), and the half spaces H_j', each defined by one of the inequalities (4.3). Hence there must be a one-to-one correspondence between these two families of half spaces.

Since, by assumption, both the z's and the x's are arranged in (anticlockwise) cyclic order so as to satisfy (4.1), we can conclude that there is a unique $0 \leqslant p < 2n$ such that

$$z^{(p+k)} = \frac{Q(x^{(k+1)} - x^{(k)})}{\langle x^{(k)}, Qx^{(k+1)} \rangle} \qquad \text{for} \quad k = 1, \ldots, 2n. \tag{4.8}$$

Since the dual norm of μ_A coincides with the max norm $\mu_{A'}$ afforded by $A' = \{x^{(1)}, \ldots, x^{(n)}\}$, and since, repeating the previous argument and renumbering the z's if necessary, $\{z^{(1)}, \ldots, z^{(2n)}\}$ constitutes the set of extreme points of the unit ball of $\mu_{A'}$, we can conclude that the extreme points $x^{(k)}$ of μ_A are given as follows:

$$x^{(k)} = \frac{Q(z^{(k+1)} - z^{(k)})}{\langle z^{(k)}, Qz^{(k+1)} \rangle} \qquad \text{for} \quad k = 1, \ldots, 2n. \tag{4.9}$$

Summarizing, we have

PROPOSITION A. *Let A be a finite (spanning) subset of R^2, and let A_0 be the set of extreme points of* conv$\{a, -a;\ a \in A\}$. *Let $A_0 = \{z^{(1)}, \ldots, z^{(2n)}\}$ be numbered so that* (4.1) *is satisfied. Then the unit ball of μ_A has $2n$ extreme points, $x^{(1)}, \ldots, x^{(2n)}$, given by* (4.9), *that is,*

$$x^{(k)} = \frac{Q(z^{(k+1)} - z^{(k)})}{\langle z^{(k)}, Qz^{(k+1)} \rangle} \qquad \textit{for} \quad k = 1, \ldots, 2n.$$

Let

$$\alpha_k = \langle z^{(k)}, Qx^{(k+1)} \rangle \qquad \text{for} \quad k = 1, \ldots, 2n. \tag{4.10}$$

Since the square of Q is $-I$, by (4.9) we have

$$\alpha_K Qx^{(k)} = z^{(k)} - z^{(k+1)} \qquad \text{for} \quad k = 1, \ldots, 2n. \tag{4.11}$$

Using this fact and the relations given in (4.1) and (4.7), we have for each $i = 1, \ldots n$

$$\begin{aligned}
\sum_{k=1}^{n} \alpha_k |\langle Qx^{(k)}, x^{(i)} \rangle| &= \sum_{k=0}^{n-1} \alpha_{i+k} |\langle Qx^{(i+k)}, x^{(i)} \rangle| \\
&= \sum_{k=0}^{n-1} \alpha_{i+k} \langle Qx^{(i+k)}, x^{(i)} \rangle \\
&= \left\langle \sum_{k=0}^{n-1} \left(z^{(i+k)} - z^{(i+k+1)} \right), x^{(i)} \right\rangle \\
&= \langle z^{(i)} - z^{(i+n)}, x^{(i)} \rangle = 2.
\end{aligned}$$

Thus

$$\sum_{k=1}^{n} \alpha_k |\langle Qx^{(k)}, x^{(i)} \rangle| = 2 \qquad \text{for all} \quad i = 1, \ldots, n. \tag{4.12}$$

Now let $B = \{w^{(1)}, \ldots, w^{(n)}\}$ be a (spanning) subset of R^2, and consider the *sum norm* afforded by B,

$$\sigma_B(x) = \sum_{i=1}^{n} |\langle w^{(i)}, x \rangle| \qquad \text{for} \quad x \in R^2. \tag{4.13}$$

We claim that the vectors given by

$$v^{(i)} = \frac{Qw^{(i)}}{\sigma_B(Qw^{(i)})},$$

$$v^{(n+i)} = -v^{(i)} \quad \text{for} \quad i = 1, \ldots, n \tag{4.14}$$

constitute the extreme points of the unit ball of σ_B. Note that these $2n$ points are all different if and only if B does not contain any two vectors that are multiples of each other. We may assume this, and in addition, letting $w^{(n+i)} = -w^{(i)}$ for $i = 1, \ldots, n$, and by suitable renumbering, we may assume that the vectors $w^{(1)}, \ldots, w^{(2n)}$ obey the relations (4.1). Then this is also true for $v^{(1)}, \ldots, v^{(2n)}$, since the v's arise from the w's by a common rotation and division by some positive real number.

Let E denote the unit ball of σ_B, and suppose that $v^{(k)}$ $(1 \leqslant k \leqslant 2n)$ is a convex combination $v^{(k)} = \tau p + (1 - \tau)q$, where $p, q \in E$, both from the boundary of E. Then

$$\begin{aligned}
\sigma_B(v^{(k)}) &= \sum_{i=1}^{n} |\langle w^{(i)}, v^{(k)}\rangle| = \sum_{i=1}^{n} |\tau\langle w^{(i)}, p\rangle + (1 - \tau)\langle w^{(i)}, q\rangle| \\
&\leqslant \tau \sum_{i=1}^{n} |\langle w^{(i)}, p\rangle| + (1 - \tau) \sum_{i=1}^{n} |\langle w^{(i)}, q\rangle| \\
&= \tau\sigma_B(p) + (1 - \tau)\sigma_B(q) = 1.
\end{aligned}$$

Equality holds if and only if there is no $1 \leqslant i \leqslant n$ such that the two scalar products $\langle w^{(i)}, p\rangle$, $\langle w^{(i)}, q\rangle$ have different algebraic signs. However, by the definition of $v^{(k)}$,

$$\begin{aligned}
0 = \langle w^{(k)}, v^{(k)}\rangle &= \langle w^{(k)}, \tau p + (1 - \tau)q\rangle \\
&= \tau\langle w^{(k)}, p\rangle + (1 - \tau)\langle w^{(k)}, q\rangle,
\end{aligned}$$

which can only be true if $\langle w^{(k)}, p\rangle$ and $\langle w^{(k)}, q\rangle$ have different signs or both vanish. Since we have to exclude the first case, both p and q are orthogonal to $w^{(k)}$, and hence multiples of each other. But then $\sigma_B(p) = \sigma_B(q) = \sigma_B(v^{(k)}) = 1$ implies $p = q = v^{(k)}$, that is, $v^{(k)}$ is an extreme point.

We still have to show that every extreme point is among the $v^{(k)}$, $k = 1, \dots, 2n$. Suppose that $z \in E$ is an additional extreme point. Then there are two consecutive vectors $v^{(j)}, v^{(j+1)}$ $(1 \leqslant j < 2n)$ such that z is a positive linear combination $z = \alpha v^{(j)} + \beta v^{(j+1)}$ with $0 < \alpha, \beta \in R$. But then, by the same reasoning as above,

$$\sigma_B(z) \leqslant \alpha \sum_{i=1}^{n} |\langle w^{(i)}, v^{(j)} \rangle| + \beta \sum_{i=1}^{n} |\langle w^{(i)}, v^{(j+1)} \rangle|$$

$$= \alpha \sigma_B(v^{(j)}) + \beta \sigma_B(v^{(j+1)}) = \alpha + \beta.$$

Equality holds if and only if there is no $1 \leqslant i \leqslant n$ such that the scalar products $\langle w^{(i)}, v^{(j)} \rangle$ and $\langle w^{(i)}, v^{(j+1)} \rangle$ have different signs. This is in fact true. Suppose, on the contrary, that for some $1 \leqslant i \leqslant n$ we have $\langle w^{(i)}, v^{(j)} \rangle = \tau > 0$, $\langle w^{(i)}, v^{(j+1)} \rangle = -\delta < 0$. Then

$$\langle w^{(i)}, \delta v^{(j)} + \tau v^{(j+1)} \rangle = \delta\tau - \tau\delta = 0.$$

From this we conclude that an extreme point, $v^{(i)} = Qw^{(i)}/\sigma_B(Qw^{(i)})$, is situated between $v^{(j)}$ and $v^{(j+1)}$, contrary to our assumption.

Completing the proof, $1 = \sigma_B(z) = \alpha + \beta$ implies that z belongs to the line segment joining $v^{(j)}$ and $v^{(j+1)}$ and therefore cannot be an extreme point.

Summarizing, we have

PROPOSITION B. *Let B be a finite (spanning) subset of R^2, and let $B' = \{w^{(1)}, \dots, w^{(n)}\}$ be the subset generated from B by elimination of multiples according to $\{\alpha x, \beta x\} \to (|\alpha| + |\beta|)x$. Then the unit ball of the sum norm σ_B afforded by B has $2n$ extreme points, $v^{(1)}, \dots, v^{(2n)}$, given by* (4.14), *that is,*

$$v^{(i)} = \frac{Qw^{(i)}}{\sigma_B(Qw^{(i)})}, \quad v^{(n+i)} = -v^{(i)} \qquad \textit{for} \quad i = 1, \dots, n.$$

Letting now

$$w^{(i)} = \frac{\alpha_i Qx^{(i)}}{2} = \frac{z^{(i)} - z^{(i+1)}}{4} \qquad \text{for} \quad i = 1, \dots, n \tag{4.15}$$

and $B = \{w^{(1)}, \ldots, w^{(n)}\}$, we consider the sum norm afforded by B,

$$\sigma_B(x) = \sum_{i=1}^{n} |\langle w^{(i)}, x\rangle| \qquad \text{for} \quad x \in R^2. \tag{4.16}$$

Then the extreme points of the unit ball of σ_B are given by

$$\pm \frac{Qw^{(i)}}{\sigma_B(Qw^{(i)})} \qquad (i = 1, \ldots, n). \tag{4.17}$$

Since by (4.12)

$$\frac{Qw^{(i)}}{\sigma_B(Qw^{(i)})} = -x^{(i)} \qquad \text{for} \quad i = 1, \ldots, n, \tag{4.18}$$

we can conclude that the unit ball of the max norm μ_A (4.2) and that of the sum norm σ_B (4.16) have the same set of extreme points. Hence $\mu_A(x)$ and $\sigma_B(x)$ are identical for all $x \in R^2$.

This proves the first part of Lemma 4.

For the second part, let a finite (spanning) subset B of R^2 be given. Then the unit ball of the sum norm σ_B afforded by B is a polyhedron with at most $2|B|$ corners. Similarly to the proof of the first part, the set of extreme points can be transformed bijectively—essentially by passing from corners to edges—into a subset A of R^2 affording a max norm μ_A that coincides with σ_B. This proves the second part of Lemma 4.

Up to this point we have proved Lemma 4 for the case where μ_A and σ_B are norms, that is, when A and B are spanning subsets of R^2. If A or B only span a 1-dimensional subspace of R^2, both μ_A and σ_B reduce to seminorms of the type $|\langle z, \cdot\rangle|$. In this case Lemma 4 is trivially valid.

As a final remark, Lemma 4 is restricted to R^2 essentially because in higher dimensions a polyhedron need not have equally many corners and faces. But even without requiring $|A| = |B|$, in R^3 it is not possible to find for any finite subset A some finite subset B such that $\mu_A = \sigma_B$, while the converse is always possible. This is related to a theorem proved in [3] stating that, within the convex cone of seminorms of R^2, the proper seminorms constitute the extremal rays. This is no longer true in R^n for $n > 2$.

The authors are indebted to Professor Andreas Dress for many stimulating remarks and helpful suggestions. Particular thanks are due to Professor Ando, who, acting as a referee of this paper, gave invaluable advice on how to reorganize the first draft to make it suitable for an LAA publication.

REFERENCES

1 P. N. Alberti and A. Uhlmann, *Dissipative Motion in State Spaces*, Teubner-Texte Math., Leipzig, 1981.
2 T. Ando, Majorization, doubly stochastic matrices, and comparison of eigenvalues, *Linear Algebra Appl.* 118:164–248 (1989).
3 A. Dress, E. Ruch, and E. Triesch, Extremals of the cone of norms and the cohomology of c-stratifications, Preprint 35, Univ. Bielefeld, 1990.
4 G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, 2nd ed., Cambridge U. P., New York, 1973.
5 K. Jacobs, Extremalpunkte konvexer Mengen, in *Selecta Math. III*, Heidelberger Taschenbücher, Springer-Verlag, New York, 1971.
6 L. W. Kantorowitsch and G. P. Akilow, *Funktionalanalyse in Normierten Räumen*, Verlag Harry Deutsch, Thun, 1978.
7 L. Nachbin, A Theorem of the Hahn-Banach type for linear transformations, *Trans. Amer. Math. Soc.* 68:28–46 (1950).
8 E. Ruch, S. Schranner, and T. H. Seligman, Generalization of a theorem of Hardy, Littlewood and Polya, *J. Math. Anal. Appl.* 76:222–229 (1980).
9 S. Sherman, A correction to "On a conjecture concerning doubly stochastic matrices," *Proc. Amer. Math. Soc.* 5:998–999 (1954).

Received 13 June 1991; final manuscript accepted 15 April 1992